
Convex Hulls of Algebraic Sets

João Gouveia¹ and Rekha Thomas²

¹ Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA, and CMUC, Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal jgouveia@math.washington.edu

² Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, USA thomas@math.washington.edu

Summary. This article describes a method to compute successive convex approximations of the convex hull of a set of points in \mathbb{R}^n that are the solutions to a system of polynomial equations over the reals. The method relies on sums of squares of polynomials and the dual theory of moment matrices. The main feature of the technique is that all computations are done modulo the ideal generated by the polynomials defining the set to the convexified. This work was motivated by questions raised by Lovász concerning extensions of the theta body of a graph to arbitrary real algebraic varieties, and hence the relaxations described here are called theta bodies. The convexification process can be seen as an incarnation of Lasserre's hierarchy of convex relaxations of a semialgebraic set in \mathbb{R}^n . When the defining ideal is real radical the results become especially nice. We provide several examples of the method and discuss convergence issues. Finite convergence, especially after the first step of the method, can be described explicitly for finite point sets.

1 Introduction

An important concern in optimization is the complete or partial knowledge of the convex hull of the set of feasible solutions to an optimization problem. Computing convex hulls is in general a difficult task, and a classical example is the construction of the integer hull of a polyhedron which drives many algorithms in integer programming. In this article we describe a method to convexify (at least approximately), an algebraic set using semidefinite programming.

By an algebraic set we mean a subset $S \subseteq \mathbb{R}^n$ described by a finite list of polynomial equations of the form $f_1(\mathbf{x}) = f_2(\mathbf{x}) = \dots = f_t(\mathbf{x}) = 0$ where $f_i(\mathbf{x})$ is an element of $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \dots, x_n]$, the polynomial ring in n variables over the reals. The input to our algorithm is the ideal generated by f_1, \dots, f_t , denoted as $I = \langle f_1, \dots, f_t \rangle$, which is the set $\{\sum_{i=1}^t g_i f_i : g_i \in \mathbb{R}[\mathbf{x}]\}$. An ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is a group under addition and is closed under multiplication by elements of $\mathbb{R}[\mathbf{x}]$. Given an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$, its *real variety*, $\mathcal{V}_{\mathbb{R}}(I) = \{\mathbf{x} \in \mathbb{R}^n :$

$f(\mathbf{x}) = 0 \forall f \in I$ is an example of an algebraic set. Given I , we describe a method to produce a nested sequence of convex relaxations of the closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$, the convex hull of $\mathcal{V}_{\mathbb{R}}(I)$, called the *theta bodies* of I . The k -th theta body $\text{TH}_k(I)$ is obtained as the projection of a *spectrahedron* (the feasible region of a semidefinite program), and

$$\text{TH}_1(I) \supseteq \text{TH}_2(I) \supseteq \cdots \supseteq \text{TH}_k(I) \supseteq \text{TH}_{k+1}(I) \supseteq \cdots \supseteq \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I))).$$

Of special interest to us are *real radical ideals*. We define the real radical of an ideal I , denoted as $\sqrt[\mathbb{R}]{I}$, to be the set of all polynomials $f \in \mathbb{R}[\mathbf{x}]$ such that $f^{2m} + \sum g_i^2 \in I$ for some $g_i \in \mathbb{R}[\mathbf{x}]$ and $m \in \mathbb{N}$. We say that I is a real radical ideal if $I = \sqrt[\mathbb{R}]{I}$. Given a set $S \subseteq \mathbb{R}^n$, the *vanishing ideal* of S , denoted as $\mathcal{I}(S)$, is the set of all polynomials in $\mathbb{R}[\mathbf{x}]$ that vanish on S . The *Real Nullstellensatz* (see Theorem 2) says that for any ideal I , $\mathcal{I}(\mathcal{V}_{\mathbb{R}}(I)) = \sqrt[\mathbb{R}]{I}$.

The construction of theta bodies for arbitrary ideals was motivated by a problem posed by Lovász. In [17] Lovász constructed the theta body of a graph, a convex relaxation of the stable set polytope of a graph which was shown later to have a description in terms of semidefinite programming. An important result in this context is that the theta body of a graph coincides with the stable set polytope of the graph if and only if the graph is perfect. Lovász observed that the theta body of a graph could be described in terms of sums of squares of real polynomials modulo the ideal of polynomials that vanish on the incidence vectors of stable sets. This observation naturally suggests the definition of a theta body for any ideal in $\mathbb{R}[\mathbf{x}]$. In fact, an easy extension of his observation leads to a hierarchy of theta bodies for all ideals as above. In [18, Problem 8.3], Lovász asked to characterize all ideals that have the property that their first theta body coincides with $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$, which was the starting point of our work. For defining ideals of finite point sets we answer this question in Section 4.

This article is organized as follows. In Section 2 we define theta bodies of an ideal in $\mathbb{R}[\mathbf{x}]$ in terms of sums of squares polynomials. For a general ideal I , we get that $\text{TH}_k(I)$ contains the closure of the projection of a spectrahedron which is described via combinatorial moment matrices from I . When the ideal I is real radical, we show that $\text{TH}_k(I)$ coincides with the closure of the projected spectrahedron, and when I is the defining ideal of a set of points in $\{0, 1\}^n$, the closure is not needed. We establish a general relationship between the theta body sequence of an ideal I and that of its real radical ideal $\sqrt[\mathbb{R}]{I}$.

Section 3 gives two examples of the construction described in Section 2. As our first example, we look at the stable sets in a graph and describe the hierarchy of theta bodies that result. The first member of the hierarchy is Lovász's theta body of a graph. This hierarchy converges to the stable set polytope in finitely many steps as is always the case when we start with a finite set of real points. The second example is a cardioid in the plane in which case the algebraic set that is being convexified is infinite.

In Section 4 we discuss convergence issues for the theta body sequence. When $\mathcal{V}_{\mathbb{R}}(I)$ is compact, the theta body sequence is guaranteed to converge to

the closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ asymptotically. We prove that when $\mathcal{V}_{\mathbb{R}}(I)$ is finite, $\text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ for some finite k . In the case of finite convergence, it is useful to know the specific value of k for which $\text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. This notion is called exactness and we characterize exactness for the first theta body when the set to be convexified is finite. There are examples in which the theta body sequence does not converge to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$. While a full understanding of when convergence occurs is still elusive, we describe one obstruction to finite convergence in terms of certain types of singularities of $\mathcal{V}_{\mathbb{R}}(I)$.

The last section gives more examples of theta bodies and their exactness. In particular we consider cuts in a graph and polytopes coming from the graph isomorphism question.

The core of this paper is based on results from [5] and [3] which are presented here with a greater emphasis on geometry, avoiding some of the algebraic language in the original results. Theorems 5, 7 and their corollaries are new while Theorem 9 is from [4]. The application of theta bodies to polytopes that arise in the graph isomorphism question is taken from [15].

Acknowledgments. Both authors were partially supported by the NSF Focused Research Group grant DMS-0757371. J. Gouveia was also partially supported by Fundação para a Ciência e Tecnologia and R.R. Thomas by the Robert R. and Elaine K. Phelps Endowed Professorship.

2 Theta bodies of polynomial ideals

To describe the convex hull of an algebraic set, we start with a simple observation about any convex hull. Given a set $S \subseteq \mathbb{R}^n$, $\text{cl}(\text{conv}(S))$, the closure of $\text{conv}(S)$, is the intersection of all closed half-spaces containing S :

$$\text{cl}(\text{conv}(S)) = \{\mathbf{p} \in \mathbb{R}^n : l(\mathbf{p}) \geq 0 \ \forall \ l \in \mathbb{R}[\mathbf{x}]_1 \text{ s.t. } l|_S \geq 0\}.$$

From a computational point of view, this observation is useless, as the right hand side is hopelessly cumbersome. However, if S is the zero set of an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$, we can define nice relaxations of the above intersection of infinitely many half-spaces using a classical strengthening of the nonnegativity condition $l|_S \geq 0$. We describe these relaxations in this section. In Section 2.1 we introduce our method for arbitrary ideals in $\mathbb{R}[\mathbf{x}]$. In Section 2.2 we specialize to real radical ideals, which occur frequently in applications, and show that in this case, much stronger results hold than for general ideals.

2.1 General ideals

Recall that given an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$, two polynomials f and g are defined to be congruent modulo I , written as $f \equiv g \pmod{I}$, if $f - g \in I$. The relation \equiv is an equivalence relation on $\mathbb{R}[\mathbf{x}]$ and the equivalence class of a polynomial f

is denoted as $f + I$. The set of all congruence classes of polynomials modulo I is denoted as $\mathbb{R}[\mathbf{x}]/I$ and this set is both a ring and a \mathbb{R} -vector space: given $f, g \in \mathbb{R}[\mathbf{x}]$ and $\lambda \in \mathbb{R}$, $(f + I) + (g + I) = (f + g) + I$, $\lambda(f + I) = \lambda f + I$, and $(f + I)(g + I) = fg + I$. Note that if $f \equiv g \pmod{I}$, then $f(\mathbf{s}) = g(\mathbf{s})$ for all $\mathbf{s} \in \mathcal{V}_{\mathbb{R}}(I)$.

We will say that a polynomial $h \in \mathbb{R}[\mathbf{x}]$ is a sum of squares (*sos*) modulo I if there exist polynomials $g_1, \dots, g_r \in \mathbb{R}[\mathbf{x}]$ such that $h \equiv \sum_{i=1}^r g_i^2 \pmod{I}$. If h is sos modulo I then we immediately have that h is nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$. In practice, it is important to control the degree of the g_i in the sos representation of h , so we will say that h is k -sos mod I if $g_1, \dots, g_r \in \mathbb{R}[\mathbf{x}]_k$, where $\mathbb{R}[\mathbf{x}]_k$ is the set of polynomials in $\mathbb{R}[\mathbf{x}]$ of degree at most k . The set of polynomials that are k -sos mod I , considered as a subset of $\mathbb{R}[\mathbf{x}]_{2k}/I$ will be denoted as $\Sigma_k(I)$.

Definition 1. Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be a polynomial ideal. We define the k -th theta body of I to be the set

$$\text{TH}_k(I) := \{\mathbf{p} \in \mathbb{R}^n : l(\mathbf{p}) \geq 0 \text{ for all } l \in \mathbb{R}[\mathbf{x}]_1 \text{ s.t. } l \text{ is } k\text{-sos mod } I\}.$$

Since, if l is sos mod I then $l \geq 0$ on $\mathcal{V}_{\mathbb{R}}(I)$, and $\text{TH}_k(I)$ is closed, $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I))) \subseteq \text{TH}_k(I)$. Also, $\text{TH}_k(I) \subseteq \text{TH}_{k-1}(I)$, since as k increases, we are potentially intersecting more half-spaces. Thus, the theta bodies of I create a nested sequence of closed convex relaxations of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

We now present a related semidefinite programming relaxation of $\mathcal{V}_{\mathbb{R}}(I)$ using the theory of moments.

For $I \subseteq \mathbb{R}[\mathbf{x}]$ an ideal, let $\mathcal{B} = \{f_0 + I, f_1 + I, \dots\}$ be a basis for the \mathbb{R} -vector space $\mathbb{R}[\mathbf{x}]/I$. We assume that the polynomials f_i representing the elements of \mathcal{B} are minimal degree representatives of their equivalence classes $f_i + I$. This makes the set $\mathcal{B}_k := \{f_i + I \in \mathcal{B} : \deg(f_i) \leq k\}$ well-defined. In this paper we will restrict ourselves to a special type of basis \mathcal{B} .

Definition 2. Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. A basis $\mathcal{B} = \{f_0 + I, f_1 + I, \dots\}$ of $\mathbb{R}[\mathbf{x}]/I$ is a θ -basis if it satisfies the following conditions:

1. $f_0(\mathbf{x}) = 1$;
2. $f_i(\mathbf{x}) = x_i$, for $i = 1, \dots, n$;
3. f_i is monomial for all i ;
4. if $\deg(f_i), \deg(f_j) \leq k$ then $f_i f_j + I$ is in the real span of \mathcal{B}_{2k} .

We will also always assume that \mathcal{B} is ordered and that $\deg(f_i) \leq \deg(f_{i+1})$.

Using Gröbner bases theory, one can see that if $\mathcal{V}_{\mathbb{R}}(I)$ is not contained in any proper affine space, then $\mathbb{R}[\mathbf{x}]/I$ always has a θ -basis. For instance, take the f_i in \mathcal{B} to be the *standard monomials* of an *initial ideal* of I with respect to some *total degree monomial ordering* on $\mathbb{R}[\mathbf{x}]$ (see for example [2]). The methods we describe below work with non monomial bases of $\mathbb{R}[\mathbf{x}]/I$ as explained in [5]. We restrict to a θ -basis in this survey for ease of exposition and since the main applications we will discuss only need this type of basis.

Fix a θ -basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I$ and define $\mathbf{f}_k[\mathbf{x}]$ to be the column vector formed by all the elements of \mathcal{B}_k in order. Then $\mathbf{f}_k[\mathbf{x}]\mathbf{f}_k[\mathbf{x}]^T$ is a square matrix indexed by \mathcal{B}_k with (i, j) -entry equal to $f_i f_j + I$. By hypothesis, the entries of $\mathbf{f}_k[\mathbf{x}]\mathbf{f}_k[\mathbf{x}]^T$ lie in the \mathbb{R} -span of \mathcal{B}_{2k} . Let $\{\lambda_{i,j}^l\}$ be the set of real numbers such that $f_i f_j + I = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l f_l + I$. We now linearize $\mathbf{f}_k[\mathbf{x}]\mathbf{f}_k[\mathbf{x}]^T$ by replacing each element of \mathcal{B}_{2k} by a new variable.

Definition 3. Let I , \mathcal{B} and $\{\lambda_{i,j}^l\}$ be as above. Let \mathbf{y} be a real vector indexed by \mathcal{B}_{2k} . The k -th combinatorial moment matrix $M_{\mathcal{B}_k}(\mathbf{y})$ of I is the real matrix indexed by \mathcal{B}_k whose (i, j) -entry is $[M_{\mathcal{B}_k}(\mathbf{y})]_{i,j} = \sum_{f_l + I \in \mathcal{B}_{2k}} \lambda_{i,j}^l y_l$.

Example 1. Let $I \subseteq \mathbb{R}[x_1, x_2]$ be the ideal $I = \langle x_1^2 - 2x_2 + x_1, x_1 x_2 \rangle$. Then a θ -basis for I would be $\mathcal{B} = \{1, x_1, x_2, x_2^2, x_2^3, x_2^4, \dots\}$. Let us construct the matrix $M_{\mathcal{B}_1}(\mathbf{y})$. Consider the vector $\mathbf{f}_1[\mathbf{x}] = (1 \ x_1 \ x_2)^T$, then

$$\mathbf{f}_1[\mathbf{x}]\mathbf{f}_1[\mathbf{x}]^T = \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{pmatrix} \equiv \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & 2x_2 - x_1 & 0 \\ x_2 & 0 & x_2^2 \end{pmatrix} \pmod{I}.$$

We now linearize the resulting matrix using $\mathbf{y} = (y_0, y_1, y_2, \dots)$, where y_i indexes the i th element of \mathcal{B} , and get

$$M_{\mathcal{B}_1}(\mathbf{y}) = \begin{pmatrix} y_0 & y_1 & y_2 \\ y_1 & 2y_2 - y_1 & 0 \\ y_2 & 0 & y_3 \end{pmatrix}.$$

The matrix $M_{\mathcal{B}_k}(\mathbf{y})$ will allow us to define a relaxation of $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ that will essentially be the theta body $\text{TH}_k(I)$.

Definition 4. Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal and \mathcal{B} a θ -basis of $\mathbb{R}[\mathbf{x}]/I$. Then

$$Q_{\mathcal{B}_k}(I) := \pi_{\mathbb{R}^n}(\{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : y_0 = 1, M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0\})$$

where $\pi_{\mathbb{R}^n}$ is projection of $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}}$ to (y_1, \dots, y_n) , its coordinates indexed by $x_1 + I, \dots, x_n + I$.

The set $Q_{\mathcal{B}_k}(I)$ is a relaxation of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. Pick $\mathbf{s} \in \mathcal{V}_{\mathbb{R}}(I)$ and define $\mathbf{y}^{\mathbf{s}} := (f_i(\mathbf{s}) : f_i + I \in \mathcal{B}_{2k})$. Then $\mathbf{y}^{\mathbf{s}}(\mathbf{y}^{\mathbf{s}})^t = M_{\mathcal{B}_k}(\mathbf{y}^{\mathbf{s}})$, $y_0^{\mathbf{s}} = 1$ and $\pi_{\mathbb{R}^n}(\mathbf{y}^{\mathbf{s}}) = \mathbf{s}$. We now show the connection between $Q_{\mathcal{B}_k}(I)$ and $\text{TH}_k(I)$.

Theorem 1. For any ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ and any θ -basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I$, we get $\text{cl}(Q_{\mathcal{B}_k}(I)) \subseteq \text{TH}_k(I)$.

Proof: We start with a general observation concerning k -sos polynomials. Suppose $h \equiv \sum_{i=1}^r g_i^2 \pmod{I}$ where $g_i \in \mathbb{R}[\mathbf{x}]_k$. Each g_i can be identified with a real row vector \hat{g}_i such that $g_i(\mathbf{x}) \equiv \hat{g}_i \mathbf{f}_k[\mathbf{x}] \pmod{I}$, and so $g_i^2 \equiv \mathbf{f}_k[\mathbf{x}]^T \hat{g}_i^T \hat{g}_i \mathbf{f}_k[\mathbf{x}]$. Denoting by P_h the positive semidefinite matrix $\sum_{i=1}^r \hat{g}_i^T \hat{g}_i$ we get $h(\mathbf{x}) \equiv$

$\mathbf{f}_k[\mathbf{x}]^T P_h \mathbf{f}_k[\mathbf{x}] \bmod I$. In general, P_h is not unique. Let \hat{h} be the real row vector such that $h(\mathbf{x}) \equiv \hat{h} \mathbf{f}_{2k}[\mathbf{x}] \bmod I$. Then check that for any column vector $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}}$, $\hat{h}\mathbf{y} = P_h \cdot M_{\mathcal{B}_k}(\mathbf{y})$, where \cdot stands for the usual entry-wise inner product of matrices.

Suppose $\mathbf{p} \in Q_{\mathcal{B}_k}(I)$, and $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}}$ such that $y_0 = 1$, $M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0$ and $\pi_{\mathbb{R}^n}(\mathbf{y}) = \mathbf{p}$. Since $\text{TH}_k(I)$ is closed, we just have to show that for any $h = a_0 + \sum_{i=1}^n a_i x_i \in \mathbb{R}[\mathbf{x}]_1$ that is k -sos modulo I , $h(\mathbf{p}) \geq 0$. Since $y_0 = 1$ and $\pi_{\mathbb{R}^n}(\mathbf{y}) = \mathbf{p}$, $h(\mathbf{p}) = a_0 + \sum_{i=1}^n a_i p_i = a_0 y_0 + \sum_{i=1}^n a_i y_i = \hat{h}\mathbf{y} = P_h \cdot M_{\mathcal{B}_k}(\mathbf{y}) \geq 0$ since $P_h \succeq 0$ and $M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0$. \square

In the next subsection we will see that when I is a real radical ideal, $\text{cl}(Q_{\mathcal{B}_k}(I))$ coincides with $\text{TH}_k(I)$.

The idea of computing approximations of the convex hull of a semialgebraic set in \mathbb{R}^n via the theory of moments and the dual theory of sums of squares polynomials is due to Lasserre [7, 8, 9] and Parrilo [20, 21]. In his original set up, the moment relaxations obtained are described via moment matrices that rely explicitly on the polynomials defining the semialgebraic set. In [8], the focus is on semialgebraic subsets of $\{0, 1\}^n$ where the equations $x_i^2 - x_i$ are used to simplify computations. This idea was generalized in [12] to arbitrary real algebraic varieties and studied in detail for zero-dimensional ideals. Laurent showed that the moment matrices needed in the approximations of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ could be computed modulo the ideal defining the variety. This greatly reduces the size of the matrices needed, and removes the dependence of the computation on the specific presentation of the ideal in terms of generators. The construction of the set $Q_{\mathcal{B}_k}(I)$ is taken from [12]. Since an algebraic set is also a semialgebraic set (defined by equations), we could apply Lasserre's method to $\mathcal{V}_{\mathbb{R}}(I)$ to get a sequence of approximations $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. The results are essentially the same as theta bodies if the generators of I are picked carefully. However, by restricting ourselves to real varieties, instead of allowing inequalities, and concentrating on the sum of squares description of theta bodies, as opposed to the moment matrix approach, we can prove some interesting theoretical results that are not covered by the general theory for Lasserre relaxations. Many of the usual results for Lasserre relaxations rely on the existence of a non-empty interior for the semialgebraic set to be convexified which is never the case for a real variety, or compactness of the semialgebraic set which we do not want to impose.

2.2 Real radical ideals

Recall from the introduction that given an ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ its real radical is the ideal

$$\sqrt[\mathbb{R}]{I} = \left\{ f \in \mathbb{R}[\mathbf{x}] : f^{2m} + \sum g_i^2 \in I, m \in \mathbb{N}, g_i \in \mathbb{R}[\mathbf{x}] \right\}.$$

The importance of this ideal arises from the Real Nullstellensatz.

Theorem 2 (Bounded degree Real Nullstellensatz [16]). *Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be an ideal. Then there exists a function $\Gamma : \mathbb{N} \rightarrow \mathbb{N}$ such that, for all polynomials $f \in \mathbb{R}[\mathbf{x}]$ of degree at most d that vanish on $\mathcal{V}_{\mathbb{R}}(I)$, $f^{2^m} + \sum g_i^2 \in I$ for some polynomials g_i such that $\deg(g_i)$ and m are all bounded above by $\Gamma(d)$. In particular $\mathcal{I}(\mathcal{V}_{\mathbb{R}}(I)) = \sqrt[\mathbb{R}]{I}$.*

When I is a real radical ideal, the sums of squares approach and the moment approach for theta bodies of I coincide, and we get a stronger version of Theorem 1.

Theorem 3. *For any $I \subseteq \mathbb{R}[\mathbf{x}]$ real radical and any θ -basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I$, $\text{cl}(Q_{\mathcal{B}_k}(I)) = \text{TH}_k(I)$.*

Proof: By Theorem 1 we just have to show that $\text{TH}_k(I) \subseteq \text{cl}(Q_{\mathcal{B}_k}(I))$. By [22, Prop 2.6], the set $\Sigma_k(I)$ of elements of $\mathbb{R}[\mathbf{x}]_{2k}/I$ that are k -sos modulo I , is closed when I is real radical. Let $f \in \mathbb{R}[\mathbf{x}]_1$ be nonnegative on $\text{cl}(Q_{\mathcal{B}_k}(I))$ and suppose $f + I \notin \Sigma_k(I)$. By the separation theorem, we can find $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_k}$ such that $\hat{f}\mathbf{y} < 0$ but $\hat{g}\mathbf{y} \geq 0$ for all $g + I \in \Sigma_k(I)$, or equivalently, $P_g \cdot M_{\mathcal{B}_k}(\mathbf{y}) \geq 0$. Since P_g runs over all possible positive semidefinite matrices of size $|\mathcal{B}_k|$, and the cone of positive semidefinite matrices of a fixed size is self-dual, we have $M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0$. Let r be any real number and consider $g_r + I := (f + r)^2 + I \in \Sigma_k(I)$. Then $\hat{g}_r\mathbf{y} = \hat{f}^2\mathbf{y} + 2r\hat{f}\mathbf{y} + r^2y_0 \geq 0$. Since $\hat{f}\mathbf{y} < 0$ and r can be arbitrarily large, y_0 is forced to be positive. So we can scale \mathbf{y} to have $y_0 = 1$, so that $\pi_{\mathbb{R}^n}(\mathbf{y}) \in Q_{\mathcal{B}_k}(I)$. By hypothesis we then have $f(\pi_{\mathbb{R}^n}(\mathbf{y})) \geq 0$, but by the linearity of f , $f(\pi_{\mathbb{R}^n}(\mathbf{y})) = \hat{f}\mathbf{y} < 0$ which is a contradiction, so f must be k -sos modulo I . This implies that any linear inequality valid for $\text{cl}(Q_{\mathcal{B}_k}(I))$ is valid for $\text{TH}_k(I)$, which proves $\text{TH}_k(I) \subseteq \text{cl}(Q_{\mathcal{B}_k}(I))$. \square

We now have two ways of looking at the relaxations $\text{TH}_k(I)$ — one by a characterization of the linear inequalities that hold on them and the other by a characterization of the points in them. These two perspectives complement each other. The inequality version is useful to prove (or disprove) convergence of theta bodies to $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ while the description via semidefinite programming is essential for practical computations. All the applications we consider use real radical ideals in which case the two descriptions of $\text{TH}_k(I)$ coincide up to closure. In some cases, as we now show, the closure can be omitted.

Theorem 4. *Let $I \subseteq \mathbb{R}[\mathbf{x}]$ be a real radical ideal and k be a positive integer. If there exists some linear polynomial g that is k -sos modulo I with a representing matrix P_g that is positive definite, then $Q_{\mathcal{B}_k}(I)$ is closed and equals $\text{TH}_k(I)$.*

Proof: For this proof we will use a standard result from convex analysis: Let V and W be finite dimensional vector spaces, $H \subseteq W$ be a cone and $A : V \rightarrow W$ be a linear map such that $A(V) \cap \text{int}(H) \neq \emptyset$. Then $(A^{-1}H)^* = A'(H^*)$ where A' is the adjoint operator to A . In particular, $A'(H^*)$ is closed in V' , the

dual vector space to V . The proof of this result follows, for example, from Corollary 3.3.13 in [1] by setting $K = V$.

Throughout the proof we will identify $\mathbb{R}[\mathbf{x}]_l/I$, for all l , with the space $\mathbb{R}^{\mathcal{B}_l}$ by simply considering the coordinates in the basis \mathcal{B}_l . Consider the inclusion map $A : \mathbb{R}^{\mathcal{B}_1} \cong \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{\mathcal{B}_{2k}}$, and let H be the cone in $\mathbb{R}^{\mathcal{B}_{2k}}$ of polynomials that can be written as a sum of squares of polynomials of degree at most k . The interior of this cone is precisely the set of sums of squares g with a positive definite representing matrix P_g . Our hypothesis then simply states that $A(\mathbb{R}^{\mathcal{B}_1}) \cap \text{int}(H) \neq \emptyset$ which implies by the above result that $A'(H^*)$ is closed. Note that H^* is the set of elements \mathbf{y} in $\mathbb{R}^{\mathcal{B}_{2k}}$ such that $\hat{h}\mathbf{y}$ is nonnegative for all $h \in H$ and this is the same as demanding $P_h \cdot M_{\mathcal{B}_k}(\mathbf{y}) \geq 0$ for all positive semidefinite matrices P_h , which is equivalent to demanding that $M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0$. So $A'(H^*)$ is just the set $\pi_{\mathbb{R}^{n+1}}(\{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : M_{\mathcal{B}_k}(\mathbf{y}) \succeq 0\})$ and by intersecting it with the plane $\{\mathbf{y} \in \mathbb{R}^{\mathcal{B}_{2k}} : y_0 = 1\}$ we get $Q_{\mathcal{B}_k}(I)$ which is therefore closed. \square

One very important case where the conditions of Theorem 4 holds is when I is the vanishing ideal of a set of 0/1 points. This is precisely the case of most interest in combinatorial optimization.

Corollary 1. *If $S \subseteq \{0, 1\}^n$ and $I = \mathcal{I}(S)$, then $Q_{\mathcal{B}_k}(I) = \text{TH}_k(I)$.*

Proof: It is enough to show that there is a linear polynomial $g \in \mathbb{R}[\mathbf{x}]$ such that $g \equiv \mathbf{f}_k[\mathbf{x}]^T A \mathbf{f}_k[\mathbf{x}] \bmod I$ for a *positive definite* matrix A and some θ -basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I$.

Let A be the matrix

$$A = \begin{pmatrix} l+1 & \mathbf{c}^t \\ \mathbf{c} & D \end{pmatrix},$$

where $l+1 = |\mathcal{B}_k|$, $\mathbf{c} \in \mathbb{R}^l$ is the vector with all entries equal to -2 , and $D \in \mathbb{R}^{l \times l}$ is the diagonal matrix with all diagonal entries equal to 4 . This matrix is positive definite since D is positive definite and its Schur complement $(l+1) - \mathbf{c}^t D^{-1} \mathbf{c} = 1$ is positive.

Since $x_i^2 \equiv x_i \bmod I$ for $i = 1, \dots, n$ and \mathcal{B} is a monomial basis, for any $f + I \in \mathcal{B}$, $f \equiv f^2 \bmod I$. Therefore, the constant (linear polynomial) $l+1 \equiv \mathbf{f}_k[\mathbf{x}]^T A \mathbf{f}_k[\mathbf{x}] \bmod I$. \square

The assumption that I is real radical seems very strong. However, we now establish a relationship between the theta bodies of an ideal and those of its real radical, showing that $\sqrt[\mathbb{R}]{I}$ determines the asymptotic behavior of $\text{TH}_k(I)$. We start by proving a small technical lemma.

Lemma 1. *Given an ideal I and a polynomial f of degree d such that $-f^{2m} \in \Sigma_k(I)$ for some $m, k \in \mathbb{N}$, the polynomial $f + \varepsilon \in \Sigma_{k+4dm}(I)$ for every $\varepsilon > 0$.*

Proof: First, note that for any $l \geq m$ and any $\xi > 0$ we have

$$f^l + \xi = \frac{1}{\xi} \left(\frac{f^l}{2} + \xi \right)^2 + \frac{1}{4\xi} (-f^{2m}) f^{2l-2m}$$

and so $f^l + \xi$ is $(dl + k)$ -sos modulo I . For $\sigma > 0$, define the polynomial $p(x)$ to be the truncation of the Taylor series of $\sqrt{\sigma^2 + 4\sigma x}$ at degree $2m - 1$ i.e.,

$$p(x) = \sum_{n=0}^{2m-1} (-1)^n \frac{(2n)!}{(n!)^2 (1-2n) \sigma^{n-1}} x^n.$$

When we square $p(x)$ we get a polynomial whose terms of degree at most $2m - 1$ are exactly the first $2m - 1$ terms of $\sigma^2 + 4\sigma x$, and by checking the signs of each of the coefficients of $p(x)$ we can see that the remaining terms of $p(x)^2$ will be negative for odd powers and positive for even powers. Composing p with f we get

$$(p(f(\mathbf{x})))^2 = \sigma^2 + 4\sigma f(\mathbf{x}) + \sum_{i=0}^{m-1} a_i f(\mathbf{x})^{2m+2i} - \sum_{i=0}^{m-2} b_i f(\mathbf{x})^{2m+2i+1}$$

where the a_i and b_i are positive numbers. In particular

$$\sigma^2 + 4\sigma f(\mathbf{x}) = p(f(\mathbf{x}))^2 + \sum_{i=0}^{m-1} a_i f(\mathbf{x})^{2i} (-f(\mathbf{x})^{2m}) + \sum_{i=0}^{m-2} b_i f(\mathbf{x})^{2m+2i+1}.$$

On the right hand side of this equality the only term who is not immediately a sum of squares is the last one, but by the above remark, since $2m + 2i + 1 > m$, by adding an arbitrarily small positive number, it becomes $(d(2m + 2i + 1) + k)$ -sos modulo I . By checking the degrees throughout the sum, one can see that for any $\xi > 0$, $\sigma^2 + 4\sigma f(\mathbf{x}) + \xi$ is $(4dm + k)$ -sos modulo I . Since σ and ξ are arbitrary positive numbers we get the desired result. \square

Lemma 1, together with the Real Nullstellensatz, gives us an important relationship between the theta body hierarchy of an ideal and that of its real radical.

Theorem 5. *Fix an ideal I . Then, there exists a function $\Psi : \mathbb{N} \rightarrow \mathbb{N}$ such that $\text{TH}_{\Psi(k)} \subseteq \text{TH}_k(\sqrt[k]{I})$ for all k .*

Proof: Fix some k , and let $f(\mathbf{x})$ be a linear polynomial that is k -sos modulo $\sqrt[k]{I}$. This means that there exists some sum of squares $s(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{2k}$ such that $f - s \in \sqrt[k]{I}$. Therefore, by the Real Nullstellensatz (Theorem 2), $-(f - s)^{2m} \in \Sigma_l(I)$ for $l, m \leq \Gamma(2k)$, where $\Gamma : \mathbb{N} \rightarrow \mathbb{N}$ depends only on the ideal I . By Lemma 1 it follows that $f - s + \varepsilon$ is $\Gamma(2k) + 8k\Gamma(2k)$ -sos modulo I for every $\varepsilon > 0$. Let $\Psi(k) := \Gamma(2k) + 8k\Gamma(2k)$. Then we have that $f + \varepsilon$ is $\Psi(k)$ -sos modulo I for all $\varepsilon > 0$. This means that for every $\varepsilon > 0$, the inequality

$f + \varepsilon \geq 0$ is valid on $\text{TH}_{\Psi(k)}(I)$ for f linear and k -sos modulo $\sqrt[k]{I}$. Therefore, $f \geq 0$ is also valid on $\text{TH}_{\Psi(k)}(I)$, and hence, $\text{TH}_{\Psi(k)}(I) \subseteq \text{TH}_k(\sqrt[k]{I})$. \square

Note that the function Ψ whose existence we just proved, can be notoriously bad in practice, as it can be much higher than necessary. The best theoretical bounds on Ψ come from quantifier elimination and so increase very fast. However, if we are only interested in convergence of the theta body sequence, as is often the case, Theorem 5 tells us that we might as well assume that our ideals are real radical.

3 Computing theta bodies

In this section we illustrate the computation of theta bodies on two examples. In the first example, $\mathcal{V}_{\mathbb{R}}(I)$ is finite and hence $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is a polytope, while in the second example $\mathcal{V}_{\mathbb{R}}(I)$ is infinite. Convex approximations of polytopes via linear or semidefinite programming have received a great deal of attention in combinatorial optimization where the typical problem is to maximize a linear function $\mathbf{c} \cdot \mathbf{x}$ as \mathbf{x} varies over the characteristic vectors $\chi^T \in \{0, 1\}^n$ of some combinatorial objects T . Since this discrete optimization problem is equivalent to the linear program in which one maximizes $\mathbf{c} \cdot \mathbf{x}$ over all $\mathbf{x} \in \text{conv}(\{\chi^T\})$ and $\text{conv}(\{\chi^T\})$ is usually unavailable, one resorts to approximations of this polytope over which one can optimize in a reasonable way. A combinatorial optimization problem that has been tested heavily in this context is the *maximum stable set problem* in a graph which we use as our first example. In [17], Lovász constructed the *theta body of a graph* which was the first example of a semidefinite programming relaxation of a combinatorial optimization problem. The hierarchy of theta bodies for an arbitrary polynomial ideal are a natural generalization of Lovász's theta body for a graph, which explains their name. Our work on theta bodies was initiated by two problems that were posed by Lovász in [18, Problems 8.3 and 8.4] suggesting this generalization and its properties.

3.1 The maximum stable set problem

Let $G = ([n], E)$ be an undirected graph with vertex set $[n] = \{1, \dots, n\}$ and edge set E . A *stable set* in G is a set $U \subseteq [n]$ such that for all $i, j \in U$, $\{i, j\} \notin E$. The maximum stable set problem seeks the stable set of largest cardinality in G , the size of which is the *stability number* of G , denoted as $\alpha(G)$.

The maximum stable set problem can be modeled as follows. For each stable set $U \subseteq [n]$, let $\chi^U \in \{0, 1\}^n$ be its *characteristic vector* defined as $(\chi^U)_i = 1$ if $i \in U$ and $(\chi^U)_i = 0$ otherwise. Let $S_G \subseteq \{0, 1\}^n$ be the set of characteristic vectors of all stable sets in G . Then $\text{STAB}(G) := \text{conv}(S_G)$ is called the *stable set polytope* of G and the maximum stable set problem is,

in theory, the linear program $\max\{\sum_{i=1}^n x_i : \mathbf{x} \in \text{STAB}(G)\}$ with optimal value $\alpha(G)$. However, $\text{STAB}(G)$ is not known apriori, and so one resorts to relaxations of it over which one can optimize $\sum_{i=1}^n x_i$.

In order to compute theta bodies for this example, we first need to view S_G as the real variety of an ideal. The natural ideal to take is, $\mathcal{I}(S_G)$, the vanishing ideal of S_G . It can be checked that this real radical ideal is

$$I_G := \langle x_i^2 - x_i \forall i \in [n], x_i x_j \forall \{i, j\} \in E \rangle \subset \mathbb{R}[x_1, \dots, x_n].$$

For $U \subseteq [n]$, let $\mathbf{x}^U := \prod_{i \in U} x_i$. From the generators of I_G it follows that if $f \in \mathbb{R}[\mathbf{x}]$, then $f \equiv g \pmod{I_G}$ where g is in the \mathbb{R} -span of the set of monomials $\{\mathbf{x}^U : U \text{ is a stable set in } G\}$. Check that

$$\mathcal{B} := \{\mathbf{x}^U + I_G : U \text{ stable set in } G\}$$

is a θ -basis of $\mathbb{R}[\mathbf{x}]/I_G$ containing $1 + I_G, x_1 + I_G, \dots, x_n + I_G$. This implies that $\mathcal{B}_k = \{\mathbf{x}^U + I_G : U \text{ stable set in } G, |U| \leq k\}$, and for $\mathbf{x}^{U_i} + I_G, \mathbf{x}^{U_j} + I_G \in \mathcal{B}_k$, their product is $\mathbf{x}^{U_i \cup U_j} + I_G$ which is $0 + I_G$ if $U_i \cup U_j$ is not a stable set in G . This product formula allows us to compute $M_{\mathcal{B}_k}(\mathbf{y})$ where we index the element $\mathbf{x}^U + I_G \in \mathcal{B}_k$ by the set U . Since $S_G \subseteq \{0, 1\}^n$ and $I_G = \mathcal{I}(S_G)$ is real radical, by Corollary 1, we have that

$$\text{TH}_k(I_G) = \left\{ \mathbf{y} \in \mathbb{R}^n : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{B}_k| \times |\mathcal{B}_k|} \text{ such that} \\ M_{\emptyset\emptyset} = 1, \\ M_{\emptyset\{i\}} = M_{\{i\}\emptyset} = M_{\{i\}\{i\}} = y_i \\ M_{UU'} = 0 \text{ if } U \cup U' \text{ is not stable in } G \\ M_{UU'} = M_{WW'} \text{ if } U \cup U' = W \cup W' \end{array} \right\}.$$

In particular, indexing the one element stable sets by the vertices of G ,

$$\text{TH}_1(I_G) = \left\{ \mathbf{y} \in \mathbb{R}^n : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{(n+1) \times (n+1)} \text{ such that} \\ M_{00} = 1, \\ M_{0i} = M_{i0} = M_{ii} = y_i \quad \forall i \in [n] \\ M_{ij} = 0 \quad \forall \{i, j\} \in E \end{array} \right\}.$$

Example 2. Let $G = (\{1, 2, 3, 4, 5\}, E)$ be a pentagon. Then

$$I_G = \langle x_i^2 - x_i \quad \forall i = 1, \dots, 5, x_1 x_2, x_2 x_3, x_3 x_4, x_4 x_5, x_1 x_5 \rangle$$

and

$$\mathcal{B} = \{1, x_1, x_2, x_3, x_4, x_5, x_1 x_3, x_1 x_4, x_2 x_4, x_2 x_5, x_3 x_5\} + I_G.$$

Let $\mathbf{y} \in \mathbb{R}^{11}$ be a vector whose coordinates are indexed by the elements of \mathcal{B} in the given order. Then

$$\text{TH}_1(I_G) = \left\{ \mathbf{y} \in \mathbb{R}^5 : \exists y_6, \dots, y_{10} \text{ s.t. } \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 \\ y_1 & y_1 & 0 & y_6 & y_7 & 0 \\ y_2 & 0 & y_2 & 0 & y_8 & y_9 \\ y_3 & y_6 & 0 & y_3 & 0 & y_{10} \\ y_4 & y_7 & y_8 & 0 & y_4 & 0 \\ y_5 & 0 & y_9 & y_{10} & 0 & y_5 \end{pmatrix} \succeq 0 \right\},$$

and $\text{TH}_2(I_G) =$

$$\left\{ \mathbf{y} \in \mathbb{R}^5 : \exists y_6, \dots, y_{10} \text{ s.t. } \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} \\ y_1 & y_1 & 0 & y_6 & y_7 & 0 & y_6 & y_7 & 0 & 0 & 0 \\ y_2 & 0 & y_2 & 0 & y_8 & y_9 & 0 & 0 & y_8 & y_9 & 0 \\ y_3 & y_6 & 0 & y_3 & 0 & y_{10} & y_6 & 0 & 0 & 0 & y_{10} \\ y_4 & y_7 & y_8 & 0 & y_4 & 0 & 0 & y_7 & y_8 & 0 & 0 \\ y_5 & 0 & y_9 & y_{10} & 0 & y_5 & 0 & 0 & 0 & y_9 & y_{10} \\ y_6 & y_6 & 0 & y_6 & 0 & 0 & y_6 & 0 & 0 & 0 & 0 \\ y_7 & y_7 & 0 & 0 & y_7 & 0 & 0 & y_7 & 0 & 0 & 0 \\ y_8 & 0 & y_8 & 0 & y_8 & 0 & 0 & 0 & y_8 & 0 & 0 \\ y_9 & 0 & y_9 & 0 & 0 & y_9 & 0 & 0 & 0 & y_9 & 0 \\ y_{10} & 0 & 0 & y_{10} & 0 & y_{10} & 0 & 0 & 0 & 0 & y_{10} \end{pmatrix} \succeq 0 \right\}.$$

It will follow from Proposition 1 below that

$$\text{STAB}(G) = \text{TH}_2(I_G) \subsetneq \text{TH}_1(I_G).$$

In general, it is a non-trivial task to decide whether two convex bodies coincide and thus to check whether a given theta body, $\text{TH}_k(I)$, equals $\overline{\text{conv}(\mathcal{V}_{\mathbb{R}}(I))}$. One technique is to show that all linear functions $l(\mathbf{x})$ such that $l(\mathbf{x}) \geq 0$ on $\mathcal{V}_{\mathbb{R}}(I)$ are k -sos mod I . Two illustrations of this method appear shortly.

The first theta body of I_G is exactly the theta body, $\text{TH}(G)$, of G as defined by Lovász [19, Lemma 2.17]. The higher theta bodies $\text{TH}_k(I_G)$ shown above give a nested sequence of convex relaxations of $\text{STAB}(G)$ extending Lovász's $\text{TH}(G)$. The problem

$$\max \left\{ \sum_{i=1}^n x_i : \mathbf{x} \in \text{TH}(G) \right\}$$

can be solved to arbitrary precision in polynomial time in the size of G via semidefinite programming. The optimal value of this semidefinite program is called the *theta number* of G and provides an upper bound on $\alpha(G)$. See [6, Chap. 9] and [26] for more on the stable set problem and $\text{TH}(G)$. The body $\text{TH}(G)$ was the first example of a semidefinite programming relaxation of a discrete optimization problem and snowballed the use of semidefinite programming in combinatorial optimization. See [14, 18] for surveys.

Recall that a graph G is *perfect* if and only if G has no induced odd cycles of length at least five, or their complements. Lovász showed that $\text{STAB}(G) = \text{TH}(G)$ if and only if G is perfect. This equality shows that the maximum stable set problem can be solved in polynomial time in the size of G when G is a perfect graph, and this geometric proof is the only one known for this complexity result. Since a pentagon is not perfect, it follows that for G a pentagon, $\text{STAB}(G) \subsetneq \text{TH}_1(I_G)$. We will see in Corollary 3 that when $\mathcal{V}_{\mathbb{R}}(I)$ is finite, then there is some finite k for which $\text{TH}_k(I) = \overline{\text{conv}(\mathcal{V}_{\mathbb{R}}(I))}$.

Since no monomial in the basis \mathcal{B} of $\mathbb{R}[\mathbf{x}]/I_G$ has degree larger than $\alpha(G)$, for any G , $\text{STAB}(G) = \text{TH}_{\alpha(G)}(I_G)$ which proves that when G is a pentagon, $\text{STAB}(G) = \text{TH}_2(I_G)$. We now prove a second, general reason why $\text{STAB}(G) = \text{TH}_2(I_G)$ when G is a pentagon.

A simple class of linear inequalities that are valid on $\text{STAB}(G)$ are the *odd cycle inequalities*, $\sum_{i \in C} x_i \leq \frac{|C|-1}{2}$, where C is an odd cycle in G .

Proposition 1. *For any graph G , all odd cycle inequalities are valid on $\text{TH}_2(I_G)$.*

Proof: Let $n = 2k+1$ and C be an n -cycle. Then $I_C = \langle x_i^2 - x_i, x_i x_{i+1} \mid i \in [n] \rangle$ where $x_{n+1} = x_1$. Therefore, $(1 - x_i)^2 \equiv 1 - x_i$ and $(1 - x_i - x_{i+1})^2 \equiv 1 - x_i - x_{i+1} \pmod{I_C}$. This implies that, mod I_C ,

$$p_i^2 := ((1 - x_1)(1 - x_{2i} - x_{2i+1}))^2 \equiv p_i = 1 - x_1 - x_{2i} - x_{2i+1} + x_1 x_{2i} + x_1 x_{2i+1}.$$

Summing over $i = 1, \dots, k$, we get

$$\sum_{i=1}^k p_i^2 \equiv k - kx_1 - \sum_{i=2}^{2k+1} x_i + \sum_{i=3}^{2k} x_1 x_i \pmod{I_C}$$

since $x_1 x_2$ and $x_1 x_{2k+1}$ lie in I_C . Define $g_i := x_1(1 - x_{2i+1} - x_{2i+2})$. Then $g_i^2 - g_i \in I_C$ and mod I_C we get that

$$\sum_{i=1}^{k-1} g_i^2 \equiv (k-1)x_1 - \sum_{i=3}^{2k} x_1 x_i, \text{ which implies } \sum_{i=1}^k p_i^2 + \sum_{i=1}^{k-1} g_i^2 \equiv k - \sum_{i=1}^{2k+1} x_i.$$

Thus, $k - \sum_{i=1}^{2k+1} x_i$ is 2-sos mod I_C .

Now let G be any graph and C be an induced $(2k+1)$ -cycle in G . Then since $I_C \subseteq I_G$, $k - \sum_{i=1}^{2k+1} x_i$ is 2-sos mod I_G which proves the result. \square

Corollary 2. *If G is an odd cycle of length at least five, $\text{STAB}(G) = \text{TH}_2(I_G)$.*

Proof: It is known that the facet inequalities of $\text{STAB}(G)$ for G an $n = 2k+1$ -cycle are:

$$x_i \geq 0 \quad \forall i \in [n], \quad 1 - x_i - x_{i+1} \geq 0 \quad \forall i \in [n], \quad \text{and} \quad k - \sum_{i=1}^n x_i \geq 0$$

(see for instance, [25, Corollary 65.12a]). Clearly, x_i is 2-sos mod I_G for all $i \in [n]$, and check that $1 - x_i - x_j \equiv (1 - x_i - x_j)^2 \pmod{I_G}$ for all $\{i, j\} \in E$. By Proposition 1, $k - \sum_{i=1}^n x_i$ is 2-sos mod I_G . If $l(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_1$ such that $l(\mathbf{x}) \geq 0$ on $\text{STAB}(G)$, then $l(\mathbf{x})$ is a nonnegative linear combination of the facet inequalities of $\text{STAB}(G)$, and hence $l(\mathbf{x})$ is 2-sos mod I_G . \square

Using the same method as in Proposition 1, one can show that the *odd antihole* and *odd wheel* inequalities [6, Chap. 9] that are valid for $\text{STAB}(G)$ are also valid for $\text{TH}_2(I_G)$. Schoenebeck [24] has shown that there is no constant k such that $\text{STAB}(G) = \text{TH}_k(I_G)$ for all graphs G (as expected, unless $\text{P}=\text{NP}$). However, no explicit family of graphs that exhibit this behavior is known.

3.2 An infinite variety

Consider the cardioid with defining equation

$$(x_1^2 + x_2^2 + 2x_1)^2 = 4(x_1^2 + x_2^2).$$

This plane curve is the real variety of the ideal $I = \langle h \rangle$ where

$$h = x_1^4 + 2x_1^2x_2^2 + x_2^4 + 4x_1^3 + 4x_1x_2^2 - 4x_2^2.$$

Points on this variety are parametrized by the angle θ made by the line segment from the origin to the point, with the x_1 -axis, by the equations:

$$x_1(\theta) = 2\cos\theta(1 - \cos\theta), \quad x_2(\theta) = 2\sin\theta(1 - \cos\theta).$$

It can be checked that the half space containing the origin defined by the tangent to the cardioid at $(x_1(\theta), x_2(\theta))$ is

$$x_1 \frac{\sin 2\theta + \sin \theta}{2\cos^2 \theta + \cos \theta - 1} - x_2 + \frac{\sin 2\theta - 2\sin \theta}{2\cos^2 \theta + \cos \theta - 1} \geq 0.$$

We now compute theta bodies of this ideal. Since the defining ideal of the cardioid is generated by a quartic polynomial, no linear polynomial is a sum of squares of linear polynomials modulo this ideal. Therefore, $\text{TH}_1(I) = \mathbb{R}^2$. A \mathbb{R} -vector space basis \mathcal{B} of $\mathbb{R}[x_1, x_2]/I$ is given by the monomials in x_1 and x_2 that are not divisible by x_1^4 . In particular,

$$\mathcal{B}_4 = \{1, x_1, x_2, x_1^2, x_1x_2, x_2^2, x_1^3, x_1^2x_2, x_1x_2^2, x_2^3, x_1^3x_2, x_1^2x_2^2, x_1x_2^3, x_2^4\} + I,$$

and suppose the coordinates of $\mathbf{y} \in \mathbb{R}^{\mathcal{B}_4}$ are indexed as follows ($y_0 = 1$):

$$\begin{array}{c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & x_1^2x_2 & x_1x_2^2 & x_2^3 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \\ \hline 1 & y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} & y_{13} \end{array}$$

To compute $M_{\mathcal{B}_2}(\mathbf{y})$ we need to express $f_i f_j + I$ as a linear combination of elements in \mathcal{B}_4 for every $f_i + I, f_j + I \in \mathcal{B}_2$ and then linearize this linear combination using \mathbf{y} . For example, $x_1^4 + I = -2x_1^2x_2^2 - x_2^4 - 4x_1^3 - 4x_1x_2^2 + 4x_2^2 + I$ which linearizes to

$$T := -2y_{11} - y_{13} - 4y_6 - 4y_8 + 4y_5.$$

Doing all these computations shows that

$$\text{TH}_2(I) = \left\{ (y_1, y_2) : \exists \mathbf{y} \in \mathbb{R}^{13} \text{ s.t. } \begin{pmatrix} 1 & y_1 & y_2 & y_3 & y_4 & y_5 \\ y_1 & y_3 & y_4 & y_6 & y_7 & y_8 \\ y_2 & y_4 & y_5 & y_7 & y_8 & y_9 \\ y_3 & y_6 & y_7 & T & y_{10} & y_{11} \\ y_4 & y_7 & y_8 & y_{10} & y_{11} & y_{12} \\ y_5 & y_8 & y_9 & y_{11} & y_{12} & y_{13} \end{pmatrix} \succeq 0 \right\}$$

For a given θ we can find the maximum $t \in \mathbb{R}$ such that $(t \cos(\theta), t \sin(\theta))$ is in $\text{TH}_2(I)$, using an SDP-solver. This point will be on the boundary of $\text{TH}_2(I)$, and if we vary θ , we will trace that boundary. In Fig. 1 we show the result obtained by repeating this procedure over 720 equally spaced directions and solving this problem using SeDuMi 1.1R3 [28].

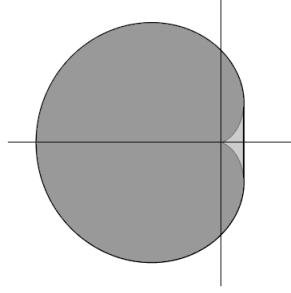


Fig. 1. $\text{TH}_2(I)$ compared with the cardioid.

From the figure, $\text{TH}_2(I)$ seems to match the convex hull of the cardioid, suggesting that $\text{TH}_2(I) = \text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. To prove this equality, we would have to construct a sos representation modulo the ideal for each of the tangents to the cardioid. Independently of that, it is clear from the figure that $\text{TH}_2(I)$ does a very good job of approximating this convex hull.

We now explain how one can also optimize a linear function over a theta body using the original sos definition of the body. For a vector $\mathbf{c} \in \mathbb{R}^n$, maximizing $\mathbf{c} \cdot \mathbf{x}$ over $\text{TH}_k(I)$ is the same as minimizing $\lambda \in \mathbb{R}$ such that $\lambda - \mathbf{c} \cdot \mathbf{x}$ is nonnegative on $\text{TH}_k(I)$. Under some mild assumptions such as compactness of $\text{TH}_k(I)$ or I being real radical, this is precisely the same as minimizing $\lambda \in \mathbb{R}$ such that $\lambda - \mathbf{c} \cdot \mathbf{x}$ is k -sos modulo I .

Consider the previous example of the cardioid and let $\mathbf{c} = (1, 1)$. If we want to optimize in that direction over $\text{TH}_2(I)$, we need to minimize λ such that $\ell_\lambda(x_1, x_2) := \lambda - x_1 - x_2$ is 2-sos modulo I . If $\mathbf{f}_2(\mathbf{x}) := (1 \ x_1 \ x_2 \ x_1^2 \ x_1 x_2 \ x_2^2)^T$, then $\ell_\lambda(x_1, x_2)$ is 2-sos modulo I if and only if there exists a 6 by 6 positive semidefinite matrix A such that

$$\ell_\lambda(x_1, x_2) \equiv \mathbf{f}_2(\mathbf{x})^T A \mathbf{f}_2(\mathbf{x}) \pmod{I}.$$

By carrying out the multiplications and doing the simplifications using the ideal one gets that this is equivalent to finding $A \succeq 0$ such that

$$\begin{aligned} \ell_\lambda(x_1, x_2) = & A_{11} + 2A_{12}x_1 + 2A_{13}x_2 + (2A_{14} + A_{22})x_1^2 \\ & + (2A_{15} + 2A_{23})x_1x_2 + (2A_{16} + A_{33} + 4A_{44})x_2^2 \\ & + (2A_{24} - 4A_{44})x_1^3 + (2A_{25} + 2A_{34})x_1^2x_2 \\ & + (2A_{26} + 2A_{35} - 4A_{44})x_1x_2^2 + 2A_{36}x_2^3 + 2A_{45}x_1^3x_2 \\ & + (2A_{46} + A_{55} - 2A_{44})x_1^2x_2^2 + 2A_{56}x_1x_2^3 + (A_{66} - A_{44})x_2^4 \end{aligned}$$

so our problem is minimizing λ such that there exists $A \succeq 0$ verifying

$$\begin{aligned} A_{11} &= \lambda; \\ A_{12} &= A_{13} = 1/2; \\ 2A_{14} + A_{22} &= 2A_{15} + 2A_{23} = 2A_{16} + A_{33} + 4A_{44} = 2A_{24} - 4A_{44} \\ &= 2A_{25} + 2A_{34} = 2A_{26} + 2A_{35} - 4A_{44} = A_{36} = A_{45} \\ &= 2A_{46} + A_{55} - 2A_{44} = A_{56} = A_{66} - A_{44} = 0 \end{aligned}$$

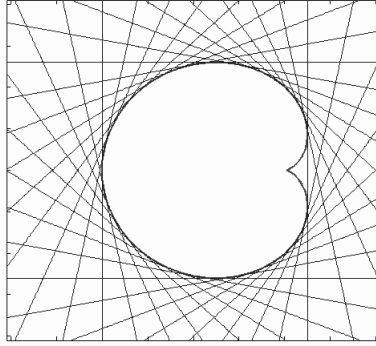


Fig. 2. Outer contour of $\text{TH}_2(I)$.

By taking $\mathbf{c} = (\cos(\theta), \sin(\theta))$, varying θ and tracing all the optimal ℓ_λ obtained one can get a contour of $\text{TH}_2(I)$. In Fig. 2 we can see all the ℓ_λ 's obtained by repeating this process 32 times using SeDuMi.

4 Convergence and Exactness

In any hierarchy of relaxations, one crucial question is that of convergence: under what conditions does the sequence $\text{TH}_k(I)$ approach $\text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$? Another important question is whether the limit is actually reached in a finite

number of steps. In this section we will give a brief overview of the known answers to these questions.

In general, convergence is not assured. In fact, for the ideal $I = \langle x^3 - y^2 \rangle$, whose real variety is a cusp, the closure of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ is the half plane defined by $x \geq 0$. However, all the theta bodies of I are just \mathbb{R}^2 . This follows immediately from the fact that no linear polynomial can ever be written as a sum of squares modulo this ideal. Fortunately, many interesting ideals behave nicely with respect to the theta body hierarchy. In particular, the central result in this area tells us that compactness of $\mathcal{V}_{\mathbb{R}}(I)$ implies convergence of the theta body hierarchy of I .

Theorem 6. *For any ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ such that $\mathcal{V}_{\mathbb{R}}(I)$ is compact,*

$$\bigcap_{k=0}^{\infty} \text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I))).$$

Proof: Let f_1, \dots, f_k be a set of generators for I . We can think of $\mathcal{V}_{\mathbb{R}}(I)$ as the compact semialgebraic set $S = \{\mathbf{x} \in \mathbb{R}^n : \pm f_i(\mathbf{x}) \geq 0\}$. Then Schmüdgen's Positivstellensatz [23] applied to S , guarantees that any linear polynomial $l(\mathbf{x})$ that is strictly positive on $\mathcal{V}_{\mathbb{R}}(I)$ has a sos representation modulo I . Hence, $l(\mathbf{x}) \geq 0$ is valid for $\text{TH}_k(I)$ for some k . \square

In general, however, we are interested in finite convergence more than in asymptotic convergence, since in that case the theta bodies can give us a representation of the closure of the convex hull of the variety as the projection of a spectrahedron, an important theoretical problem on its own. We will say that an ideal I is **TH_k-exact** if $\text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$ and **TH-exact** if it is **TH_k-exact** for some k . We first note that to study TH-exactness of an ideal depends only on the real radical of the ideal.

Theorem 7. *An ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is TH-exact if and only if $\sqrt[k]{I}$ is TH-exact.*

Proof: The “if” direction follows from Theorem 5, while the “only if” direction follows from the fact that $I \subseteq \sqrt[k]{I}$. \square

An important case in which TH-exactness holds is when $\mathcal{V}_{\mathbb{R}}(I)$ is finite.

Corollary 3. *If I is an ideal such that $\mathcal{V}_{\mathbb{R}}(I)$ is finite, then I is TH-exact.*

Proof: By Theorem 7 we can assume that I is real radical. If $\mathcal{V}_{\mathbb{R}}(I) = \{\mathbf{p}_1, \dots, \mathbf{p}_m\}$, then we can construct interpolators $g_1, \dots, g_m \in \mathbb{R}[\mathbf{x}]$ such that $g_i(\mathbf{p}_j) = 1$ if $j = i$ and 0 otherwise. Assume that k is the highest degree of a g_i . Since $g_i - g_i^2$ vanishes on $\mathcal{V}_{\mathbb{R}}(I)$, $g_i - g_i^2 \in \sqrt[k]{I} = I$ and therefore, g_i is k -sos mod I for $i = 1, \dots, m$.

If $f(\mathbf{x}) \geq 0$ on $\mathcal{V}_{\mathbb{R}}(I)$, then check that $f(\mathbf{x}) \equiv \sum_{i=1}^m f(\mathbf{p}_i)g_i(\mathbf{x}) \pmod{I}$. Since the polynomial on the right hand side is a nonnegative combination of the g_i 's which are k -sos mod I , f is k -sos mod I . In particular, all the

linear polynomials that are nonnegative on $\mathcal{V}_{\mathbb{R}}(I)$ are k -sos mod I and so, $\text{TH}_k(I) = \text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. \square

Under the stronger assumption that $\mathcal{V}_{\mathbb{C}}(I)$ is finite, TH-exactness of I follows from a result of Parrilo (see [13] for a proof). The work in [10] implies that when $\mathcal{V}_{\mathbb{R}}(I)$ is finite, the bodies $Q_{\mathcal{B}_k}(I)$ converges in a finite number of steps, a slightly weaker result than the one in Corollary 3. Finite varieties are of great importance in practice. For instance, they are precisely the feasible regions of the 0/1 integer programs.

To study TH-exactness in more detail, a better characterization is needed.

Theorem 8. *A real radical ideal $I \subseteq \mathbb{R}[\mathbf{x}]$ is TH_k -exact if and only if all linear polynomials l that are non-negative in $\mathcal{V}_{\mathbb{R}}(I)$ are k -sos modulo I .*

Proof: The “if” direction follows from the definition of theta bodies. The “only if” direction is a consequence of the proof of Theorem 3 which said that when I is real radical, $\text{TH}_t(I) = \text{cl}(Q_{\mathcal{B}_t}(I))$ for all t . Recall that in the second part of that proof we showed that any linear inequality valid for $\text{cl}(Q_{\mathcal{B}_t}(I))$ was t -sos mod I . Therefore, if $\text{TH}_k(I) = \text{cl}(\text{conv}(\mathcal{V}_{\mathbb{R}}(I)))$, then being nonnegative over $\text{cl}(Q_{\mathcal{B}_k}(I)) = \text{TH}_k(I)$ is equivalent to being nonnegative over $\mathcal{V}_{\mathbb{R}}(I)$ giving us the intended result. \square

The last general result on exactness that we will present, taken from [4], is a negative one. Given a point $\mathbf{p} \in \mathcal{V}_{\mathbb{R}}(I)$ we define the *tangent space* of \mathbf{p} , $T_{\mathbf{p}}(I)$, to be the affine space passing through \mathbf{p} and orthogonal to the vector space spanned by the gradients of all polynomials that vanish on $\mathcal{V}_{\mathbb{R}}(I)$. We say that a point \mathbf{p} in $\mathcal{V}_{\mathbb{R}}(I)$ is *convex-singular*, if it is on the boundary of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ and $T_{\mathbf{p}}(I)$ is not tangent to $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ i.e., $T_{\mathbf{p}}(I)$ intersects the relative interior of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$.

Theorem 9. *An ideal with a convex-singularity is not TH-exact.*

Proof: Let I be an ideal and \mathbf{p} a convex-singular point of $\mathcal{V}_{\mathbb{R}}(I)$, and $J = \mathcal{I}(\mathcal{V}_{\mathbb{R}}(I))$. By Theorem 7, it is enough to show that J is not TH-exact. Let $l(\mathbf{x})$ be a linear polynomial that is positive on the relative interior of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ and zero at \mathbf{p} , which we can always find since \mathbf{p} is on the boundary of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$. If J was TH_k -exact for some k , then since J is real radical, by Theorem 8 we would be able to write $l(\mathbf{x}) = \sigma(\mathbf{x}) + g(\mathbf{x})$ where σ is a sum of squares and $g \in J$. Evaluating at \mathbf{p} , we see that $\sigma(\mathbf{p}) = 0$, which since σ is a sum of squares, implies that $\nabla\sigma(\mathbf{p}) = 0$. Therefore, we must have $\nabla l = \nabla g(\mathbf{p})$. Let \mathbf{q} be a point in the relative interior of $\text{conv}(\mathcal{V}_{\mathbb{R}}(I))$ that is also in $T_{\mathbf{p}}(I)$. Then by the definition of $T_{\mathbf{p}}(I)$,

$$l(\mathbf{q}) = (\mathbf{q} - \mathbf{p})\nabla l = (\mathbf{q} - \mathbf{p})\nabla g(\mathbf{p}) = 0$$

which contradicts our choice of l . Hence, I is not TH_k -exact. \square

Even in the presence of convex-singular points, the theta bodies of the defining ideal can approximate the convex hull of the real variety arbitrarily well, but they will not converge in finitely many steps. An example of this is the ideal $I = \langle x^4 - x^3 + y^2 \rangle$. It has a compact real variety with a convex-singularity at the origin. Hence we have asymptotic convergence of the theta bodies, but not finite convergence. The first theta body of this ideal is \mathbb{R}^2 and in Fig. 3 we can see that the next two theta bodies already closely approximate the convex hull of the variety.

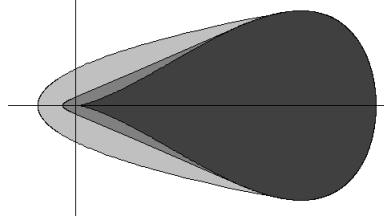


Fig. 3. The second and third theta bodies of $I = \langle x^4 - x^3 + y^2 \rangle$.

In the remainder of this section we focus on a more restricted exactness question. Problem 8.3 in [18] motivates the question of which ideals in $\mathbb{R}[\mathbf{x}]$ are TH_1 -exact. Of particular interest are vanishing ideals of finite sets of points in \mathbb{R}^n and ideals arising in combinatorial optimization.

Theorem 10. *Let S be a finite set of points in \mathbb{R}^n , then $I = \mathcal{I}(S)$ is TH_1 -exact if and only if for each facet F of the polytope $\text{conv}(S)$ there exists a hyperplane H parallel to F such that $S \subseteq F \cup H$.*

Proof: We can assume without loss of generality that $\text{conv}(S)$ is full-dimensional, otherwise we could restrict ourselves to its affine hull. Assume I is TH_1 -exact and let F be a facet of $\text{conv}(S)$. We can find a linear polynomial l such that $l(\mathbf{p}) = 0$ for all $\mathbf{p} \in F$ and $l(\mathbf{p}) \geq 0$ for all $\mathbf{p} \in S$. By Theorem 8, l must be 1-sos modulo I which implies $l \equiv \sum_{i=1}^r g_i^2 \text{ mod } I$ for some linear polynomials g_i . In particular, $l(\mathbf{p}) = \sum_{i=1}^r (g_i(\mathbf{p}))^2$ for all $\mathbf{p} \in S$, and since l vanishes in all points of $S \cap F$, so must the g_i . This implies that all the g_i 's vanish on the hyperplane $\{\mathbf{x} \in \mathbb{R}^n : l(\mathbf{x}) = 0\}$, since they are linear. This is equivalent to saying that every g_i is a scalar multiple of l , and therefore, $l \equiv \lambda l^2 \text{ mod } I$ for some nonnegative λ . Then $l - \lambda l^2 \in I$ implies that $S \subseteq \{\mathbf{x} : l(\mathbf{x}) - \lambda l(\mathbf{x})^2 = 0\}$ which is the union of the hyperplanes $\{\mathbf{x} : l(\mathbf{x}) = 0\}$ and $\{\mathbf{x} : l(\mathbf{x}) = \sqrt{1/\lambda}\}$. So take H to be the second hyperplane.

Suppose now for each facet F of $\text{conv}(S)$ there exists a hyperplane H parallel to F such that $S \subseteq F \cup H$. This implies that for any facet F and a fixed linear inequality $l_F(\mathbf{x}) \geq 0$ that is valid on S and holds at equality on F , l_F attains the same nonzero value at all points in $S \setminus F$. By scaling we

can assume that value to be one, which implies that $l_F(1 - l_F)$ is zero at all points in S , and hence, $l_F(1 - l_F) \in I$. But then $l_F = l_F^2 + l_F(1 - l_F)$ is 1-sos mod I . By Farkas Lemma, any valid linear inequality for a polytope can be written as a nonnegative linear combination of the facet inequalities of the polytope. Therefore, any linear polynomial l that is nonnegative over S is a positive sum of 1-sos polynomials mod I , and hence, l is 1-sos mod I . Now using Theorem 8, I is TH_1 -exact. \square

We call a polytope P a k -level polytope if for any facet F of P and any supporting hyperplane H of F , there are $k - 1$ hyperplanes H_1, \dots, H_{k-1} all parallel to H such that the vertices of P are contained in $H \cup H_1 \cup \dots \cup H_{k-1}$. Theorem 10 states that $\mathcal{I}(S)$ is TH_1 -exact if and only if S is the set of vertices of a 2-level polytope. Polytopes with integer vertices that are 2-level are called *compressed polytopes* in the literature [27, 29].

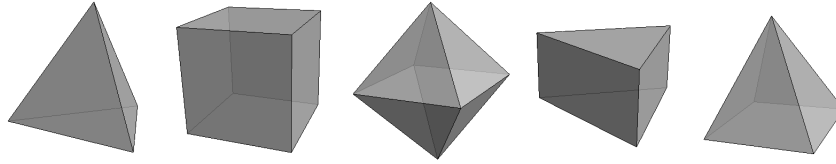


Fig. 4. The five 2-level polytopes in \mathbb{R}^3 up to affine transformations.

Examples of 2-level polytopes include simplices, cubes and cross-polytopes. In \mathbb{R}^3 , up to affine transformations, there are only five different 2-level polytopes and they are shown in Fig. 4. An example of a polytope that is not 2-level is the truncated cube shown in Figure 5. Three parallel translates of the hyperplane spanned by the slanted face are needed to contain all vertices of the polytope, and hence this is a 3-level polytope. Combinatorial properties of 2-level polytopes can be found in [5].

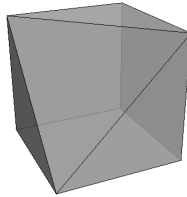


Fig. 5. This polytope is 3-level but not 2-level.

The result in Theorem 10 can be generalized to a (weak) sufficient criterion for TH_k -exactness.

Theorem 11. *If $S \subseteq \mathbb{R}^n$ is the set of vertices of a $(k+1)$ -level polytope then $\mathcal{I}(S)$ is TH_k -exact.*

Proof: As before, we may assume that $\text{conv}(S)$ is full-dimensional, and as observed in the proof of Theorem 10, it is enough to prove that for every facet F of $\text{conv}(S)$, if h_F is a linear polynomial that is zero on F and non-negative on S then h_F is k -sos modulo $\mathcal{I}(S)$. We can also assume that F is contained in the hyperplane $\{\mathbf{x} \in \mathbb{R}^n : x_1 = 0\}$, and that all points in S have nonnegative first coordinate, as all the properties we are interested in are invariant under translations.

Then, by scaling, we may assume $h_F(\mathbf{x}) = x_1$, and the $(k+1)$ -level property tells us that there exist k positive values a_1, \dots, a_k such that all points in S have the first coordinate in the set $\{0, a_1, \dots, a_k\}$. Then we can construct a one variable Lagrange interpolator g of degree k such that $g(0) = 0$ and $g(a_i) = \sqrt{a_i}$ for $i = 1, \dots, k$. This will imply that $h_F(\mathbf{x}) \equiv g(x_1)^2$ modulo $\mathcal{I}(S)$ and so h_F is k -sos modulo $\mathcal{I}(S)$. \square

This sufficient condition is very restrictive in general, and in fact can be arbitrarily bad.

Example 3. We saw in Corollary 2 that if G is a $(2k+1)$ -cycle then the ideal I_G is TH_2 -exact. However notice that we need $k+1$ parallel translates of the hyperplane $k = \sum x_i$ to cover all the vertices of $\text{STAB}(G)$, since the incidence vectors of the stable sets in G can have anywhere from 0 to k entries equal to one. Theorem 11 would only guarantee that I_G is TH_k -exact.

When $\mathcal{V}_{\mathbb{R}}(I)$ is not finite, TH_1 -exactness becomes harder to guarantee. A useful result in this direction is an alternative characterization of the first theta body of any ideal I . Given any ideal I , if we take all convex quadratic polynomials in I , and intersect the convex hulls of their zero sets, we obtain $\text{TH}_1(I)$ exactly. This result, proved in [5, Theorem 5.4], can be used in some simple cases to both prove or disprove TH_1 -exactness.

Many other questions on convergence remain open. For example, there is no known example of a smooth variety whose theta body hierarchy does not converge finitely, but there is also no reason to believe that such an example does not exist. In all studied examples of smooth hypersurfaces, convergence happens at the first non-trivial theta body. Either an explanation for this behavior, or a few examples of badly behaved smooth hypersurfaces, would be an important first step toward a better understanding of the theta body hierarchy for smooth varieties. Non-compact varieties have also not received much attention.

5 More Examples and Applications

5.1 The maximum cut problem

In section 3.1 we computed the theta body hierarchy for the maximum stable set problem. Another important problem in combinatorial optimization is the maximum cut (maxcut) problem. Given a graph $G = ([n], E)$, we say that $C \subseteq E$ is a **cut** of G , if there exists some partition of the vertices into two sets V and W such that C is the set of edges between a vertex in V and a vertex in W . The maxcut problem is the problem of finding a cut in G of maximum cardinality. Theta bodies of the maxcut problem were studied in [3], and in this subsection we will present some of those results.

The model we use is similar to the one used in the stable set problem. The characteristic vector of a cut C of G , is the vector $\chi_C \in \{-1, 1\}^E$ defined as $\chi_C(e) = -1$ if $e \in C$ and 1 otherwise. Let C_G be the collection of all characteristic vectors of cuts of G , and $\mathcal{I}(C_G)$ the vanishing ideal of C_G . The maximum cardinality cut of G can be found, in principle, by optimizing $\sum_{e \in E} x_e$ over C_G . However, this is a difficult problem and the size of the the maxcut can be approximated by optimizing $\sum_{e \in E} x_e$ over $\text{TH}_k(\mathcal{I}(C_G))$.

It is not hard to show that $\mathcal{I}(C_G)$ is generated by the polynomials $1 - x_e^2$ for all $e \in E$, together with the polynomials $1 - \prod_{e \in K} x_e$ where $K \subseteq E$ is a chordless cycle. Using this one can construct a θ -basis for $\mathcal{I}(C_G)$, but to do that we need to introduce another combinatorial concept. Given an even set $T \subseteq [n]$, we call a subgraph H of G a T -join if the set of vertices of H with odd degree is precisely T . For example an \emptyset -join is a cycle, and the minimal $\{s, t\}$ -join is the shortest path from s to t . It is clear that there exists a T -join if and only if T has an even number of nodes in each of the connected components of G . Define $\mathcal{T}_G := \{T \subseteq [n] : \exists T\text{-join in } G\}$, and for each $T \in \mathcal{T}_G$, choose H_T to be a T -join with a minimal number of edges. Define $\mathcal{T}_k := \{T \subseteq \mathcal{T}_G : |H_T| \leq k\}$, and note that $\mathcal{T}_1 = \{\emptyset\} \cup E$. Then one can show that $\mathcal{B} = \{\prod_{e \in T} x_e + \mathcal{I}(C_G) : T \in \mathcal{T}_G\}$ is a θ -basis for $\mathcal{I}(C_G)$, and we can therefore identify \mathcal{B} with \mathcal{T}_G and each \mathcal{B}_k with \mathcal{T}_k .

We can now give a description of the theta bodies of $\mathcal{I}(C_G)$.

$$\text{TH}_k(\mathcal{I}(C_G)) = \left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{|\mathcal{T}_k| \times |\mathcal{T}_k|} \text{ such that} \\ M_{e\emptyset} = y_e, \forall e \in E \\ M_{TT} = 1, \forall T \in \mathcal{T}_k \\ M_{T_1 T_1} = M_{T_3 T_4} \text{ if } T_1 \Delta T_2 = T_3 \Delta T_4 \end{array} \right\}.$$

In particular, since $\mathcal{T}_1 = \{\emptyset\} \cup E$ we get

$$\text{TH}_1(\mathcal{I}(C_G)) = \left\{ \mathbf{y} \in \mathbb{R}^E : \begin{array}{l} \exists M \succeq 0, M \in \mathbb{R}^{\{\emptyset\} \cup E \times \{\emptyset\} \cup E} \text{ such that} \\ M_{\emptyset\emptyset} = M_{ee} = 1, \forall e \in E \\ M_{\emptyset e} = M_{e\emptyset} = y_e \forall e \in E \\ M_{ef} = y_g \text{ if } \{e, f, g\} \text{ is a triangle in } G \\ M_{ef} = M_{gh} \text{ if } \{e, f, g, h\} \text{ is a square in } G \end{array} \right\}.$$

These relaxations are very closely related to some older hierarchies of relaxations for the maxcut problem. In particular they have a very interesting relation with the relaxation introduced in [11]. The theta body relaxation as it is, is not very strong, as shown by the following proposition.

Proposition 2 ([3, Corollary 5.7]). *Let G be the n -cycle. Then the smallest k for which $\mathcal{I}(C_G)$ is TH_k -exact is $k = \lceil n/4 \rceil$.*

Convergence can be sped up for graphs with large cycles by taking a chordal closure of the graph, computing the theta body for this new graph, and then projecting to the space of edges of the original graph. If instead of the chordal closure one takes the complete graph over the same vertices, we essentially recover the hierarchy introduced in [11].

One can use the usual theta body hierarchy together with Theorem 10 to solve a question posed by Lovász in [18, Problem 8.4]. Motivated by the fact that a graph is perfect if and only if the stable set ideal I_G is TH_1 -exact, Lovász asked to characterize the graphs that were “cut-perfect”, i.e., graphs G for which $\mathcal{I}(C_G)$ is TH_1 -exact. It turns out that such a “strong cut-perfect graph theorem” is not too hard to derive.

Proposition 3 ([3, Corollary 4.12]). *Given a graph G , the ideal $\mathcal{I}(C_G)$ is TH_1 -exact if and only if G had no K_5 minor and no chordless circuit of length greater than or equal to five.*

5.2 Permutation Groups

In [15, Sect. 4], De Loera et al. study the behavior of the theta body hierarchy when applied to ideals associated to permutation groups. Let $A \subseteq S_n$ be a subgroup of permutations and identify each element in A with a permutation matrix. We can then see these permutation matrices as vectors in $\mathbb{R}^{n \times n}$ and define I_A to be their vanishing ideal. One of the main goals in [15, Sect. 4] is to provide sufficient conditions for the TH_1 -exactness of I_A .

A permutation group A is *permutation summable* if for all $P_1, \dots, P_m \in A$ such that all entries of $\sum P_i - I$ are nonnegative, $\sum P_i - I$ is itself a sum of matrices in A . For example, S_n itself is permutation summable as a direct consequence of Birkhoff’s Theorem. We say that A is *strongly fixed-point free* if the only permutation in A that has fixed points is the identity. In terms of matrices this is equivalent to saying that an element of A is either the identity or has a zero diagonal. This trivially implies that all strongly fixed-point free groups are permutation summable.

Proposition 4 ([15, Theorem 4.5]). *If $A \subseteq S_n$ is a permutation summable group, then I_A is TH_1 -exact.*

A very interesting class of permutation groups is the automorphism groups of graphs. Given a graph $G = ([n], E)$, $\text{Aut}(G) \subseteq S_n$ is defined to be the group

of vertex permutations σ such that $\{i, j\} \in E$ implies $\{\sigma(i), \sigma(j)\} \in E$. The *automorphism polytope* of G , denoted as $P_{\text{Aut}(G)}$, is the convex hull of $\text{Aut}(G)$, where again we are identifying a permutation with its matrix representation. There has been some study of ways to relax this polytope. See [30] for instance. One such relaxation is the polytope P_G of all the points $P \in \mathbb{R}^{n \times n}$ such that

$$PA_G = A_GP; \quad \sum_{i=1}^n P_{i,j} = 1, \quad 1 \leq j \leq n;$$

$$\sum_{j=1}^n P_{i,j} = 1, \quad 1 \leq i \leq n; \quad P_{i,j} \geq 0, \quad 1 \leq i, j \leq n;$$

where A_G is the $n \times n$ adjacency matrix of G . If in the constraints of P_G we replace $P_{i,j} \geq 0$ by $P_{i,j} \in \{0, 1\}$, then we get precisely $\text{Aut}(G)$. In fact, $P_{\text{Aut}(G)}$ is the integer hull of P_G i.e., the convex hull of the integer points of P_G . If $P_G = P_{\text{Aut}(G)}$ then G is said to be *compact*. It is not hard to show that $\text{TH}_1(I_{\text{Aut}(G)}) \subseteq P_G$, so the first theta body provides an approximation of $P_{\text{Aut}(G)}$ that is at least as good as the linear approximation P_G , but one can actually say more.

Proposition 5. [15, Theorem 4.4] *The class of compact graphs is strictly included in the class of graphs with TH_1 -exact automorphism ideal. In particular, let G_1, \dots, G_m be k -regular graphs (all vertices have degree k) that are compact and G their disjoint union. Then $I_{\text{Aut}(G)}$ is TH_1 -exact, but G is compact if and only if $G_1 \cong \dots \cong G_m$.*

Finally note that while theta bodies of automorphism groups of graphs have interesting theoretical properties, computing with them tends to be quite hard, as there is no easy general way to obtain a good θ -basis for this problem. In fact, knowing if a graph has a non-trivial automorphism, the *graph automorphism problem*, is equivalent to knowing if a θ -basis for $I_{\text{Aut}(G)}$ has more than one element. Determining the complexity class of this problem is a major open question.

References

1. Jonathan M. Borwein and Adrian S. Lewis. *Convex analysis and nonlinear optimization. Theory and examples*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3. Springer, New York, second edition, 2006.
2. David Cox, John Little, and Donal O'Shea. *Ideals, Varieties and Algorithms*. Springer-Verlag, New York, 1992.
3. João Gouveia, Monique Laurent, Pablo Parrilo, and Rekha Thomas. A new hierarchy of semidefinite programming relaxations for cycles in binary matroids and cuts in graphs. arXiv:0907.4518.

4. João Gouveia and Tim Netzer. Positive polynomials and projections of spectrahedra. arXiv:0911.2750.
5. João Gouveia, Pablo A. Parrilo, and Rekha R. Thomas. Theta bodies for polynomial ideals. *SIAM Journal on Optimization*, 20(4):2097–2118, 2010.
6. Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, Vol. 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993.
7. Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.
8. Jean B. Lasserre. An explicit equivalent positive semidefinite program for nonlinear 0-1 programs. *SIAM J. Optim.*, 12(3):756–769, 2002.
9. Jean B. Lasserre. Convex sets with semidefinite representation. *Math. Program.*, 120:457–477, 2009.
10. Jean B. Lasserre, Monique Laurent, and Philipp Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Found. Comput. Math.*, 8(5):607–647, 2008.
11. Monique Laurent. Semidefinite relaxations for max-cut. In *The sharpest cut*, MPS/SIAM Ser. Optim., pp. 257–290. SIAM, Philadelphia, PA, 2004.
12. Monique Laurent. Semidefinite representations for finite varieties. *Math. Program.*, 109(1, Ser. A):1–26, 2007.
13. Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, Vol. 149 of *IMA Volumes in Mathematics and its Applications*. Springer, 2009.
14. Monique Laurent and Franz Rendl. Semidefinite programming and integer programming. In Karen Aardal, George Nemhauser, and Robert Weismantel, editors, *Handbook on Discrete Optimization*, pp. 393–514. Elsevier B.V., 2005.
15. Jesus De Loera, Chris Hillar, Peter N. Malkin, and Mohamed Omar. Recognizing graph theoretic properties with polynomial ideals. arXiv:1002.4435.
16. Henri Lombardi. Une borne sur les degrés pour le théorème des zéros réel effectif. In *Real algebraic geometry (Rennes, 1991)*, Vol. 1524 of *Lecture Notes in Math.*, pp. 323–345. Springer, Berlin, 1992.
17. László Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.
18. László Lovász. Semidefinite programs and combinatorial optimization. In *Recent advances in algorithms and combinatorics*, Vol. 11 of *CMS Books Math./Ouvrages Math. SMC*, pp. 137–194. Springer, New York, 2003.
19. László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optim.*, 1(2):166–190, 1991.
20. Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000.
21. Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Prog.*, 96(2, Ser. B):293–320, 2003.
22. Victoria Powers and Claus Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1(1):71–88, 2001.
23. Konrad Schmüdgen. The K -moment problem for compact semi-algebraic sets. *Math. Ann.*, 289(2):203–206, 1991.
24. Grant Schoenebeck. Linear level Lasserre lower bounds for certain k -csp. In *FOCS*, pp. 593–602. IEEE Computer Society, 2008.

25. Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, Vol. 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chaps. 39–69.
26. F. Bruce Shepherd. The theta body and imperfection. In *Perfect graphs*, Wiley-Intersci. Ser. Discrete Math. Optim., pp. 261–291. Wiley, Chichester, 2001.
27. Richard P. Stanley. Decompositions of rational convex polytopes. *Ann. Discrete Math.*, 6:333–342, 1980. Combinatorial mathematics, optimal designs and their applications (Proc. Sympos. Combin. Math. and Optimal Design, Colorado State Univ., Fort Collins, Colo., 1978).
28. Jos F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optim. Methods Softw.*, 11/12(1-4):625–653, 1999. Interior point methods.
29. Seth Sullivant. Compressed polytopes and statistical disclosure limitation. *Tohoku Math. J. (2)*, 58(3):433–445, 2006.
30. Gottfried Tinhofer. Graph Isomorphism and Theorems of Birkhoff Type. *Computing*, (36):285–300, 1986.